GRASPA 2015
BARI 15-16 JUNE

A semi–parametric approach in the estimation of the structural risk in environmental applications

R. Pappadà[1,*], E. Perrone[2], F. Durante[3] and G. Salvadori[4]

[1] *Department of Economics, Business, Mathematics and Statistics, University of Trieste, I-34127 Trieste (Italy); rpappada@units.it*
[2] *Institut für Angewandte Statistik, Johannes Kepler Universität, A-4040 Linz (Austria); elisa.perrone@jku.at*
[3] *Faculty of Economics and Management, Free University of Bozen-Bolzano, I-39100 Bolzano (Italy); fabrizio.durante@unibz.it*
[4] *Dipartimento di Matematica e Fisica, Università del Salento, I-73100 Lecce (Italy); gianfausto.salvadori@unisalento.it*
[*] *Corresponding author*

Abstract. *In environmental applications, the estimation of the structural risk is crucial. A statistical model for the behavior of the input variables is generally required, possibly accounting for different dependence structures among such variables. Copulas represent a suitable tool for dealing with natural extremes and non-linear dependencies. Two semi-parametric procedures for the approximation of, respectively, Extreme Value and Archimedean copulas, are proposed in order to provide a model for the estimation of the structural risk. The approximating techniques are evaluated by Monte Carlo tests, and illustrated via a case study concerning a preliminary rubble mound breakwater design.*

Keywords. *Copula; Monte Carlo; Return Period; Structural risk*

1 Introduction

Recent developments in environmental statistics have shown the great potential of copulas in a multivariate risk assessment framework (see, for instance, [1, 4]). In the traditional structural approach, the response $\mathbf{Y} = (Y_1, \dots, Y_m)$ of a given structure to some environmental (random) loads $\mathbf{X} = (X_1, \dots, X_d)$ is evaluated via a (multi-dimensional) structural equation, generally written as

$$\mathbf{Y} = \Psi(\mathbf{X}; \boldsymbol{\theta}), \tag{1}$$

where the structure function Ψ is known, and $\boldsymbol{\theta}$ represents a set of possible covariates. Beside the knowledge of the physical response of the structure to the loads of interest, the structural approach requires a statistical model for the behavior of the variables affecting the structure, which are generally dependent. Copulas may supply valuable dependence models accounting for a wide variety of dependence patterns. A multivariate approach based on copulas is adopted, by considering two structures often used in practice: the Extreme Value (hereafter, *EV*) copulas for dealing with rare (catastrophic) events, and the Archimedean copulas for their desirable analytical properties. The selection of an appropriate

copula model still presents some troublesome issues. As a valid alternative, this work provides two semi-parametric approximation procedures to, respectively, EV and Archimedean copulas, which can be used in the estimation of the structural risk. A "minimal" approximating strategy is adopted, involving the least number of parameters and presenting the fewest fitting difficulties.
In the following, a classical problem in coastal engineering is considered to illustrate the procedures. The problem concerns the preliminary design of a rubble mound breakwater and the target is to compute, for prescribed Return Periods (hereafter, *RP*), the weight $\mathbf{Y} = W$ of a concrete cube element forming the breakwater structure, assuming that the environmental load is given by the pair of dependent variables $\mathbf{X} = (H, D)$, where H represents the significant wave height (in meters), and D the sea storm duration (in hours). Under the assumption that $\mathbf{X}$ can be modeled by an EV or an Archimedean copula, the main idea is to provide a semi-parametric approximation of such a copula, avoiding the fit of any specific parametric model. The (highly non-linear) structure function Ψ in Eq. (1) is calibrated for the buoy of Alghero (Sardinia, Italy), as illustrated in [11], Section 3, where all the parameters (water density, ρ_W, cube density, ρ_S, number of units displaced, N_d, gravitational acceleration, g) are specified:

$$W = \rho_S \cdot \left[H \left(\frac{2\pi H}{g\,[4.597 \cdot H^{0.328}]^2} \right)^{0.1} \right]^3 \Bigg/ \left[\left(\frac{\rho_S}{\rho_W} - 1 \right) \cdot \left(1 + \frac{6.7 \cdot N_d^{0.4}}{(3600D/[4.597 \cdot H^{0.328}])^{0.3}} \right) \right]^3 . \quad (2)$$

As marginal distributions for H and D, we adopt suitable Generalized Weibull laws. Via the Monte Carlo strategy sketched in [12], suitable pairs (H, D)'s can be simulated from the approximating copula, and corresponding sample values of W can be calculated, yielding an estimate of the empirical distribution function of W. Then suitable design values w_T's for W, corresponding to specific RP T's, can be computed via the standard formula

$$w_T = F_W^{-1}(1 - \mu/T) \quad (3)$$

where F_W is the law of W, and $\mu > 0$ is the mean inter-arrival time between successive sea storms.

2 Semi-parametric approximations

In order to provide a suitable approximating copula, two characterizing functions known as Pickands dependence function and Kendall distribution function are used in the EV and Archimedean cases, respectively. We recall that a bivariate copula $\mathbf{C}$ is simply (the restriction of) a joint distribution over $\mathbf{S} = [0,1] \times [0,1]$, whose univariate marginals are Uniform. A 2-copula $\mathbf{C}$ is bivariate EV ([5]) if there exists a Pickands dependence function $\mathbf{A}$ such that $\mathbf{C}(u,v) = \exp(\ln(uv)\,\mathbf{A}(\ln v/\ln(uv)))$, for all $(u,v) \in \mathbf{S}$, being $\mathbf{A}\colon [0,1] \to [1/2,1]$ a convex function satisfying the constraint $\max\{t, 1-t\} \leq \mathbf{A}(t) \leq 1$, for all $t \in [0,1]$. The uni-dimensional function $\mathbf{A}$ uniquely identifies $\mathbf{C}$. Also, the Kendall distribution function $\mathbf{K_C}\colon [0,1] \to [0,1]$ associated with $\mathbf{C}$ is defined as $\mathbf{K_C}(t) = \mathbf{P}(\mathbf{C}(U,V) \leq t)$, where $t \in [0,1]$ is a probability level, and U, V are Uniform random variables on $[0,1]$ with 2-copula $\mathbf{C}$ ([2]). Then, the generator γ of a suitable Archimedean copula $\mathbf{C}(u,v) = \gamma^{[-1]}(\gamma(u) + \gamma(v))$ can be expressed in terms of the Kendall's function $\mathbf{K}$ associated with $\mathbf{C}$ via the formula $\gamma(t) = \exp\left(\int_{t_0}^{t} 1/(x - \mathbf{K}(x))\,dx\right)$, with $t, t_0 \in (0,1)$ and t_0 arbitrary. The equivalent differential form is provided by [8], Theorem 4.3.4. Then, $\mathbf{C}$ can be taken as the representative element of the class of copulas sharing the same function $\mathbf{K}$.
Now, the approximating technique consists in two steps: first, fit a suitable auxiliary function (the Pickands and the Kendall functions, respectively) to the available data, and, secondly, construct an appropriate copula generator exploiting such an auxiliary function. It is worth stressing that a piecewise linear interpolation scheme (both for the EV and the Archimedean case) is adopted, since it represents

a natural and basic semi-parametric approximation to the generators of interest. In addition, the convergence of such linear approximations is guaranteed under minimal conditions, which, in turn, yields the convergence of the corresponding copulas ([10, 9]).

2.1 The EV case

Let $\mathcal{X}_n$ denote a set of abscissas $x_0 = 0 < x_1 < \cdots < x_{n-1} < x_n = 1$ in $\mathbf{I} = [0,1]$, with $n \in \mathbf{N}$. Moreover, $\mathbf{X}$ will denote a sample of size $N > 0$ of i.i.d. bivariate sea storms (H,D)'s.
A possible procedure for supplying approximate structural estimates is as follows. (i) Provide an estimate $\mathbf{A}_n$ of the Pickands dependence function $\mathbf{A}$ associated with $\mathbf{X}$ at the points of $\mathcal{X}_n$ via some common rank-based estimator (available in the R-package "Copula" [7]). (ii) Generate two independent variates u and w Uniform over $(0,1)$. (iii) Numerically compute $v = c_u^{-1}(w)$ where the function $c_u(v) = \partial \mathbf{C}_n(u,v)/\partial u$ can be calculated by, first, computing the approximating copula $\mathbf{C}_n$ associated with $\mathbf{A}_n$, and then applying suitable numerical formulas for estimating the partial derivative. Finally, the pair $\left(H = F_H^{-1}(u), D = F_D^{-1}(v)\right)$ can be obtained by inversion of marginal laws and Eq. (2) used to fix the corresponding cube weight W.
Once a suitable sample of simulated cube weights $\mathbf{W}$'s is made available, Eq. (3) can be used to provide approximate design values w_T's, for prescribed RP's T's, by inverting the empirical distribution function of the $\mathbf{W}$'s. It is possible to show that $\mathbf{A}_n$ is a consistent and asymptotic Normal estimator of $\mathbf{A}$ ([3]). Thus, the corresponding approximating copula $\mathbf{C}_n$ provides a consistent estimator of the true copula $\mathbf{C}$.

2.2 The Archimedean case

Let $\mathcal{K}_n = \{k_0 = 0 \le k_1 \le \cdots \le k_{n-1} \le k_n = 1\}$ be a set of estimates of the Kendall distribution function computed at the points of $\mathcal{X}_n$. Given the set of points (x_i, k_i), with $i = 0, \ldots, n$, the procedure is as follows. (i) Generate two independent variates s and t Uniform over $(0,1)$. (ii) Numerically calculate $w = \gamma_n(\mathbf{K}_n^{-1}(t))$, where an approximation of the Archimedean generator γ_n is computed by using suitable numerical integration procedures. (iii) Numerically compute $u = \gamma_n^{-1}(s\,w)$ and $v = \gamma_n^{-1}((1-s)\,w)$.
Finally, use the marginal laws to calculate the pair $\left(H = F_H^{-1}(u), D = F_D^{-1}(v)\right)$, and the expression for W given by Eq. (2) to fix the corresponding cube weight W. Approximate design values w_T's, for prescribed RP's T's, can be obtained by inverting the empirical distribution function of the sample of simulated cube weights $\mathbf{W}$'s (Eq. (3)). It is possible to show that $\mathbf{K}_n$ is a consistent and asymptotic Normal estimator of $\mathbf{K}$ ([6]). In turn, the copula $\mathbf{C}_n$ associated with $\mathbf{K}_n$ is a consistent estimator of $\mathbf{C}$ (see also [10]).

3 Monte Carlo tests

In order to test the performance of the two approaches outlined above, the Gumbel family of copulas is chosen, since it is both EV and Archimedean. The evaluation of the performances is based on the response of the structure considered, and on the comparison of the estimates with the "true" Gumbel values that arise when the model is known. In particular, the values corresponding to different design quantiles of W are considered, corresponding to standard RP's of 10, 20, 50, 100 years. A sample of sea storms of size N is generated, and used to provide approximate design values of the cube weights. The latter ones are then compared with the "true" Gumbel ones, with Kendall's rank correlation coefficient τ. By varying N and τ, it is possible to draw fairly exhaustive statistical scenarios concerning the performance of the EV and Archimedean strategies introduced. Overall, the results (not reported here) indicate that the proposed approach may provide valuable estimates of the sought quantiles, without making recourse to any specific parametric model.

4 Conclusions

The procedures outlined in this work are essentially based on the fact that both EV and Archimedean copulas may be suitable to describe the joint statistical behavior of the variables involved in many environmental applications. One main advantage is that they can be generated by using auxiliary one-dimensional functions (the Pickands dependence function and the Kendall distribution function), and estimated using the available data. The techniques outlined in this work may provide valuable approximations to some (large) Return Period design values, commonly used in risk assessment. Monte Carlo tests and a robustness-sensitivity study (not discussed here) are carried out in order to investigate the performance of the approximating algorithms. As a result, the procedures presented in this work may provide valuable indications for a preliminary assessment of the structural risk and the choice of a suitable parametric model.

References

[1] Durante, F., Okhrin, O. (2015). Estimation procedures for exchangeable Marshall copulas with hydrological application. *Stochastic Environmental Research and Risk Assessment* **29**, 205–226.

[2] Genest, C., Rivest, LP. (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association* **88**, 1034–104.3

[3] Genest, C., Segers, J. (2009). Rank-based inference for bivariate Extreme Value copulas. *Annals of Statistics* **37**, 2990–3022.

[4] Grimaldi, S., Koutsoyiannis, D., Piccolo, D., Schumann, A. (2009). Recent developments of statistical tools for hydrological application. *Physics and Chemistry of the Earth* **34**, 595–595.

[5] Gudendorf, G., Segers, J. (2010). Extreme-value copulas. In: Jaworski, P., Durante, F., Härdle, W., Rychlik, T. (eds). *Copula Theory and its Applications, Lecture Notes in Statistics - Proceedings* **198**, 127–145. Springer. Berlin Heidelberg.

[6] Hofert, M., Mächler, M., McNeil, AJ. (2012). Likelihood inference for Archimedean copulas in high dimensions under known margins. *Journal of Multivariate Analysis* **10** 133–150.

[7] Hofert, M., Kojadinovic, I., Maechler, M., Yan, J. (2013). Copula: Multivariate Dependence with Copulas. R package version 0.999-7 edn.

[8] Nelsen, R. (2006). *An introduction to copulas*. 2nd edn. Springer-Verlag, New York.

[9] Pappadà, R. (2009). *Copule bivariate dei Valori Estremi*. Tesi di Laurea Specialistica (in Italian), Dipartimento di Matematica, Università del Salento, Lecce (Italy).

[10] Salvadori, G., Durante, F., Perrone, E. (2013). Semi-parametric approximation of the Kendall's distribution and multivariate Return Periods. *Journal de la Société Française de Statistique* **154**, 151–173.

[11] Salvadori, G., Durante, F., Tomasicchio, G.R., D'Alessandro, F. (2015). Practical guidelines for the multivariate assessment of the structural risk in coastal and off-shore engineering. *Coastal Engineering* **95**, 77–83.

[12] Straub, D. (2014). Engineering risk assessment. In: Klüppelberg, C., Straub, D., Welpe, IM. (eds) *Risk - A Multidisciplinary Introduction* 333–362. Springer International Publishing Switzerland.